ROCKETDYNE • A DIVISION OF NORTH AMERICAN AVIATION, INC.

VARIABLE MESH MULTISTEP METHODS FOR ORDINARY DIFFERENTIAL EQUATIONS

by

R. Van Wyk

Research Report 67-12

September 1967

Technical Review and Approval:

A. L. Huebner
Manager
Propulsion Physics and Mathematics
Research Division

T. A. Coultas
Manager
Propulsion Physics,
Processes and Applications
Research Division

Release Approval:

R. B. Lawhead
Manager
Physical and Engineering Sciences
Research Division

FOREWORD

This work was supported by the National Aeronautics and Space Administration Office of Advanced Research and Technology under Contract No. NAS7-470. The NASA technical manager of this contract is Dr. R. H. Wilson, Chief, Applied Mathematics Branch.

ABSTRACT

A multistep predictor-corrector method for the numerical solution of ordinary differential equations is developed. The difference equations employed are generalizations, for the case of variable mesh spacing, of previous formulas requiring fixed step size. In addition to retaining the high local accuracy and convergence properties of the earlier methods, the variable mesh method is developed in a form conducive to the generation of effective criteria for the selection of subsequent step sizes in the step by step solution of differential equations. These criteria are based on considerations of truncation error, convergence of corrector iterations, and an extensive treatment of relative numerical stability. The algorithm has been tested extensively and compared with other methods. The results of the comparison favor the new method.

CONTENTS

Foreword ii
Abstract iii
Introduction 1
Variable Mesh Multistep Formulas 3
Numerical Stability 10
Criteria for Selecting Mesh Increments 19
Numerical Testing and Comparison With Other Methods 24
References 31

ILLUSTRATIONS

1. Region of Stability and Convergence of Corrector Iterations of Variable Mesh Method 14
2. Example of Ambiguity in Relative Stability Definition . . 17
3. Regions of Relative Stability and Convergence of Corrector Iterations of Variable Mesh Method for Systems of Differential Equations 20
4. Relative Errors for Problem 1 29

TABLES

1. Problems Used to Test Variable Mesh Method 26
2. Comparison of Relative Error 28

VARIABLE MESH MULTISTEP METHODS FOR ORDINARY DIFFERENTIAL EQUATIONS*

by

R. Van Wyk

Rocketdyne
A Division of North American Aviation, Inc.
Canoga Park, California

1. INTRODUCTION. A great deal of research effort has been directed toward the numerical solution of first order nonlinear ordinary differential equations because of the practical importance of such problems. The most widely used numerical methods that have been developed for these problems provide approximate values of the solution at discrete points according to a stepwise computation beginning at an initial point for which the solution is known. These methods are called one-step methods if the calculation of the solution at a given point depends explicitly on values of the solution and one or more of its derivatives at only one previous point. Multistep methods require values at two or more previous points. One-step (Runge-Kutta) methods are very convenient because the step increments can be changed readily from step to step as desired and because the solution in the initial steps is calculated with the same formulas as used in subsequent steps. Multistep methods, although less convenient, are usually more efficient because, by making use of the calculations of more than one previous step, less computer time is required to achieve the same accuracy as achieved with a one-step method.

*This work was supported by the National Aeronautics and Space Administration Office of Advanced Research and Technology under Contract No. NAS7-470.

The research reported here was directed toward the development and testing of variable mesh multistep methods which not only preserve the efficiency due to the multistep structure but improve this efficiency by permitting as much freedom in the variation of the step increments as is afforded by one-step methods. Care was taken to formulate the basic difference equations in a manner conducive to the development of effective criteria for selecting the variable mesh increments as the calculation progresses. In the following pages the basic algorithm is described and the analysis and practical considerations justifying the mesh criteria are presented. The mesh criteria were subjected to extensive numerical testing, and in addition, the algorithm was compared with known methods in the numerical solution of selected differential equations. The results of this and other experimental work are summarized in later sections.

The problem of starting the computation, that is, the requirement of computing the solution at the first few points by a separate technique in order to initialize multistep methods, is not emphasized here for two reasons. First, because of the variable mesh formulation, the calculation is only initialized once and never has to be restarted as would be required in changing the step size while using a fixed step size, multistep method. In the second place, fairly general starting procedures are readily available for incorporation with the variable mesh method because the step increments used in the starting procedure can be smaller than those used in the subsequent calculation. For example, the starting procedure outlined in [1] for the variable mesh method consists of simply using the one-step Adams-Bashforth/Adams-Moulton formulas for the first step, the two-step formulas for the second step, and the three-step formulas for third step. The same step size is used for each of these three initial steps, and it is chosen small enough to yield the desired accuracy at the first point. There is little danger of exceeding this error at the second and third points since higher order formulas are used.

The algorithm developed in the following pages is, to a great extent, general purpose in nature, and therefore other methods or modifications of this method may be superior in specific situations. For example, systems of differential equations describing greatly different relaxation rates have received particular attention in recent years because step size is limited by the fastest rate while the total interval of integration is determined by the slowest rate. A special purpose method, such as the one due to Treanor [2], has obvious advantages for systems of this type. The writer is presently studying a modification of the variable mesh, multistep approach particularly suited to these problems. This work will be reported at a later date.

2. VARIABLE MESH MULTISTEP FORMULAS. The initial value problems of interest are represented by differential equations of the form

$$\frac{dy}{dx} = F(x,y), \tag{1}$$

with initial condition $y(x_o) = y_o$. It is assumed at the outset that F is continuous and satisfies the Lipschitz condition that guarantees the existence of a unique, continuous and differentiable solution [3]. The continuity of higher derivatives will be required later in the discussion of truncation error.

We will use the usual notation in which y_n denotes the computed value of $y(x_n)$ and y'_n denotes $F(x_n, y_n)$. It is assumed that the computed solution is obtained recursively by one or more formulas of the following type:

$$y_{n+1} = a_o y_n + a_1 y_{n-1} + a_2 y_{n-2} + a_3 y_{n-3}$$

$$+ h(b_{-1} y'_{n+1} + b_o y'_n + b_1 y'_{n-1} + b_2 y'_{n-2} + b_3 y'_{n-3}). \tag{2}$$

Here h denotes the current step size, $x_{n+1} - x_n$, and is permitted to vary with n. The coefficients a_i and b_i are also variable and it will be convenient later to express them in terms of mesh parameters α, β, and γ defined by

$$\alpha = (x_n - x_{n-1})/h,$$

$$\beta = (x_n - x_{n-2})/h, \tag{3}$$

and $$\gamma = (x_n - x_{n-3})/h.$$

We require $\gamma > \beta > \alpha > 0$.

For the case of fixed step size, α, β, and γ have the constant values 1,2, and 3, respectively. Henrici [3] has defined conditions of consistency and stability for fixed h and has shown that they are necessary and also sufficient,when taken together, for convergence of y_n to $y(x_n)$ as $h \to 0$. The stability condition requires that no root of the equation

$$\rho^4 - a_0\rho^3 - a_1\rho^2 - a_2\rho - a_3 = 0 \tag{4}$$

exceed one in modulus and roots of unit modulus must be simple. The consistency condition requires that equation (2) be exact if y(x) is either constant or linear.

An analogous equivalence theorem holds in the variable mesh case. It follows immediately that the stability and consistency conditions are necessary for convergence in the variable mesh case because they are necessary in the special case of fixed mesh. Henrici's proof of sufficiency has been generalized by the author to account for the case of

variable step size, but is omitted here because of its length. For this case the two conditions are required to hold for each different step size used in the integration.

The validity of the generalized equivalence theorem is not restricted to difference equations with only the number of terms actually shown in (2). However since we will restrict the present discussion to fourth order methods--that is, methods with error terms proportional to the fifth power of h--the terms shown are adequate. The optimum order to use in a given application depends heavily on the degree of accuracy desired, but fourth order is a reasonable compromise for medium range accuracy--say two to six significant figures. With fixed step size it is often desirable to vary the order within a given application in order to maintain a desired accuracy. The variable mesh procedure, on the other hand, has the advantage of achieving the same objective without switching from formulas of one order to those of another.

Both explicit (predictor) and implicit (corrector) variable mesh formulas are used. The explicit equation has $b_{-1} \equiv 0$ while the implicit usually can be solved by iteration. The coefficients in (2) for the two formulas are determined in part by requiring satisfaction of the stability and consistency conditions. By requiring exactness for $F(x,y) \equiv 0$, the relation, $a_0 + a_1 + a_2 + a_3 = 1$, is imposed, from which it follows that one root of equation (4) is unity. The other three roots ("parasitic"), arising because a fourth order difference equation is used in place of a first order differential equation, are all zero if we select $a_0 = 1$, $a_1 = a_2 = a_3 = 0$. Making this choice, we say that the condition of stability is satisfied optimally.

The consistency condition is satisfied by the additional requirement of exactness for $y = x-x_n$, which yields, for the predictor, $b_o + b_1 + b_2 + b_3 = 1$. Combining this with the requirements of exactness for $\frac{1}{2}(x-x_n)^2$, $\frac{1}{3}(x-x_n)^3$, $\frac{1}{4}(x-x_n)^4$, the b_i of the predictor,

$$y_{n+1} = y_n + h(b_o y'_n + b_1 y'_{n-1} + b_2 y'_{n-2} + b_3 y'_{n-3}), \tag{5}$$

are determined recursively as follows:

$$b_3 = \frac{2(2+3\alpha)(\beta+\alpha) + 3(1-2\alpha^2)}{12\gamma(\gamma-\alpha)(\beta-\gamma)}$$

$$b_2 = \frac{2+3\alpha-6\gamma(\gamma-\alpha)b_3}{6\beta(\beta-\alpha)}$$

$$b_1 = -\frac{1}{2\alpha}(1+2\gamma b_3+2\beta b_2)$$

$$b_0 = 1 - b_3 - b_2 - b_1 . \tag{6}$$

Similarly, a corrector of the form

$$y_{n+1} = y_n + h(d_{-1} y'_{n+1} + d_o y'_n + d_1 y'_{n-1} + d_2 y'_{n-2}) \tag{7}$$

is found with coefficients

$$d_2 = \frac{1 + 2\alpha}{12\,\beta(1+\beta)(\beta-\alpha)}$$

$$d_1 = -\frac{2\beta + 1}{12\alpha(1+\alpha)(\beta-\alpha)}$$

$$d_o = \frac{1}{2} - d_2(1+\beta) - d_1(1+\alpha)$$

$$d_{-1} = 1 - d_2 - d_1 - d_o \tag{8}$$

For the special case of fixed step size, the above predictor and corrector formulas reduce to the widely accepted Adams-Bashforth and Adams-Moulton formulas respectively. In this connection one is reminded of the formulas presented by Nordsieck in a paper which, like the present paper, also emphasizes the advantages of changing step size [4]. Although the algorithm of Nordsieck is substantially different from that presented here, it is similar in the sense that his basic integration formulas are equivalent to the Adams formulas. However, the formulation used by Nordsieck appears to be much less conducive to the development of effective mesh selection criteria than is the formulation presented above. This claim is corroborated by evidence obtained when both methods, complete with their respective recommended mesh selection criteria, were applied to selected differential equations. This work is described in more detail in a later section.

Assuming continuous higher derivatives of $F(x,y)$, it is evident upon comparing equation (5) with an appropriate Taylor Series representation for $y(x_{n+1})$ that the truncation error in (5) can be represented as

$$P_n \frac{h^5}{5!} y_n^v + 0(h^6),$$

where the coefficient P_n depends on α, β, and γ. If we consider the residual error resulting from the application of (5) to the polynomial $(x-x_n)^5$, we find that

$$P_n = 1 - 5\,(b_1\alpha^4+b_2\beta^4+b_3\gamma^4)$$

$$= 1 + \frac{5}{12}\,[3(\alpha+\beta+\gamma) + 4(\alpha\beta+\alpha\gamma+\beta\gamma) + 6\alpha\beta\gamma]. \qquad (9)$$

Similarly, if the error in (7) is taken in the form

$$C_n \frac{h^5}{5!} y_n^v + O(h^6),$$

C_n is found to be given by

$$C_n = 1 - 5\,(d_{-1}+d_1\alpha^4+d_2\beta^4)$$

$$= 1 - \frac{5}{12}\,(3+2\alpha\beta+\alpha+\beta). \qquad (10)$$

Various alternative modes of utilization of the predictor and corrector formulas are available in practice. For example, the predictor can be used without employing the corrector at all. On the other hand, if the corrector is used, it usually is used iteratively, with the predictor providing the first guess. Qualitatively, some of the arguments for and against the various alternatives are as follows:

a) Number of derivative evaluations per step. The "predictor-only" mode requires only one evaluation per step. If one correction is employed, a second evaluation is usually made after the correction, but is not absolutely necessary. In general, n corrections require either n or n+1 derivative evaluations, depending on whether a final evaluation is or is not carried out. Evaluations of complicated derivative functions frequently require a predominant portion of the total computer time.

b) Truncation error. Implementation of the corrector reduces the truncation error. (It is a simple exercise to show that $|C_n| < |P_n|$.)

c) Numerical stability. With regard to both absolute and relative stability, the regions of stability become less restrictive as the number of corrections is increased. Incidentally, these regions become more restrictive as order is increased.

d) Availability of mesh criteria. More effective procedures for automatically selecting the mesh increments can be developed for some modes than for others. This consideration favors a predictor-corrector mode with at least two applications of the corrector.

An empirical program was carried out whereby the various modes were compared in the actual numerical solution of selected differential equations. The mesh increments were selected in a manner such that the total number of derivative evaluations was the same for each mode. This work is not reported in detail here since an even more extensive testing program of a similar nature for the case of fixed step size was carried out and reported in detail by Hull and Creemer [5]. Their conclusions are in agreement with those reached in the present study. The results indicate the following trends:

a) It is usually wasteful to use more than two corrections per step. For example (letting p denote predictor; d derivative evaluation; and c corrector), p-d-c-d-c is more accurate than p-d-c-d-c-d-c, where 50-percent larger step sizes are used with the latter.

b) Ending each step with a derivative evaluation is not as efficient as ending with a correction. For example, p-d-c-d-c and p-d-c (with half as large step sizes) are both more accurate than p-d-c-d. By the same token p-d-c is better than p-d, thereby eliminating the "predictor-only" mode.

Based on the above results and the preceding qualitative arguments, it was concluded that the mode p-d-c-d-c is favored over such a broad range of applications that the other modes can be neglected. However, the mode p-d-c is nearly as good in many cases.

A variable mesh generalization of a predictor (applicable only in the p–d–c–d mode) due to Crane and Klopfenstein [6] was also employed in the tests but was consistently the least accurate of the methods tested.

3. NUMERICAL STABILITY. The condition of stability used to establish convergence in the previous section does not guarantee numerical stability for $h > 0$. A more appropriate analysis of numerical stability is presented here.

First note that each corrector iteration is performed according to the equation

$$c_{n+1}^{(k+1)} = y_n + hd_{-1}F(x_{n+1}, c_{n+1}^{(k)}) + h\sum_{i=0}^{2} d_i y'_{n-i},$$

where the superscript k denotes the k^{th} iteration. Subtracting this equation from (7) and employing the mean value theorem gives

$$y_{n+1} - c_{n+1}^{(k+1)} = \lambda d_{-1}(y_{n+1} - c_{n+1}^{(k)}),$$

where

$$\lambda = h\left(\frac{\partial F}{\partial y}\right)_{\substack{x=x_{n+1} \\ y=\eta}},$$

for some η between y_{n+1} and $c_{n+1}^{(k)}$. Thus the following condition is required for convergence of the corrector iterations:

$$|\lambda d_{-1}| < 1. \tag{11}$$

It is assumed that condition (11) is satisfied in the following discussion, and in fact this condition will be used in the mesh selection procedures described in the next section.

It is also assumed for the purpose of the numerical stability analysis that λ is constant, a standard assumption in the literature for fixed step size. By appropriate choice of h at each step, λ can be made nearly constant in the variable mesh case.

In practice, however, this assumption is usually violated with fixed mesh methods, even when procedures to frequently double or halve the step size are included. Furthermore when numerical stability is the controlling factor, it is good policy to keep h as large as possible without forcing λ beyond its limitation imposed by the threat of instability. Thus in this case, the mesh increments used are actually considerably suboptimal at most steps with fixed mesh methods. On the other hand, the variable mesh feature obviously allows much better optimization when the integration is stability limited. Of course when it is not stability limited, variations in λ are inconsequential.

Initially let us consider the mode which employs a prediction and k corrections with a derivative evaluation after each prediction and correction. Let ϵ_n denote the propagated error, $y(x_n) - c_n^{(k)}$. Then it can be shown that ϵ_n satisfies the difference equation

$$\epsilon_{n+1} = \epsilon_n \left[1 + \sum_{j=1}^{k} \lambda^j d_{-1}^{j-1} (d_o + d_{-1}) + \lambda^{k+1} d_{-1}^k b_o \right]$$

$$+ \sum_{i=1}^{2} \epsilon_{n-i} \sum_{j=1}^{k} \lambda^j d_{-1}^{j-1} d_i + \sum_{i=1}^{3} \epsilon_{n-i} \lambda^{k+1} d_{-1}^k b_i ,$$

except for the fifth order truncation error. The effect of the predictor on the propagated error decreases with increasing k because the factor $(\lambda d_{-1})^k$ multiplies the b_i in the above equation. In the limit the corrector alone determines the error propagation, the equation being given by

$$\epsilon_{n+1}(1-\lambda d_{-1}) - \epsilon_n(1+\lambda d_0) - \epsilon_{n-1}\lambda d_1 - \epsilon_{n-2}\lambda d_2 = 0. \qquad (12)$$

In practice, when the mesh increments are small enough to provide a reasonably small truncation error, the corrector iterations beyond the second are essentially redundant. Hence the above difference equation for the propagated error in the corrector alone also adequately represents error propagation for the recommended mode, p-d-c-d-c.

If the difference equation (12) has constant coefficients, its solution ϵ_n can be expressed in terms of the roots ρ_i of the polynomial equation

$$\rho^3(1-\lambda d_{-1}) - \rho^2(1+\lambda d_0) - \rho\lambda d_1 - \lambda d_2 = 0 \qquad (13)$$

by $\epsilon_n = k_1\rho_1^n + k_2\rho_2^n + k_3\rho_3^n$ (slightly modified in the case of a multiple root), where the k_i are constants. Equation (12) has constant coefficients as required provided the d_i are constant as well as λ. The d_i are constant in the case of fixed mesh. In the variable mesh case, it is this investigator's experience that the d_i vary very slowly when the integration is stability limited. This is due to the fact that the ratio α of mesh increments from step to step remains nearly constant, and the d_i are constant when the mesh parameters α, β, and γ are constant. (When α is constant, β and γ are the constants $\alpha + \alpha^2$ and $\alpha^2 + \alpha^3$, respectively.) Thus it is reasonable to add the assumption of constant d_i for the stability analysis, and in view of the above remarks it becomes convenient to treat numerical stability in terms of the two parameters λ and α.

When $\lambda = 0$, the fundamental root of the characteristic equation (13) is unity and the others are zero. When $\lambda \neq 0$, one or both of the latter roots may become larger in modulus than the fundamental root. This is a condition of relative numerical instability [7], whereas absolute numerical instability occurs whenever any root is greater than one in modulus or when a root of unit modulus is a multiple root. Applying these conditions as definitions, regions of both relative and absolute stability have been computed by tracking the roots of (13). These regions are shown in Fig. 1 in terms of the parameters λ and α. Although it is interesting to note the behavior for very large and small α, in practice α actually remains fairly close to unity. Also shown in Fig. 1 are the curves $\lambda d_{-1} = \pm 1$, which indicate the region for which the corrector iterations converge, and within which the stability regions have meaning.

The variable mesh formulas are applicable for systems of differential equations of the form

$$\frac{dy^{(i)}}{dx} = F_i(x, y^{(1)}, y^{(2)}, \ldots, y^{(N)}), \quad i = 1,2,\ldots,N. \tag{14}$$

In this case, equation (12) for the propagated error is replaced by

$$(I - d_{-1}hG)\bar{\epsilon}_{n+1} - (I + d_0 hG)\bar{\epsilon}_n - d_1 hG\bar{\epsilon}_{n-1} - d_2 hG\bar{\epsilon}_{n-2} = 0, \tag{15}$$

where $\bar{\epsilon}_n$ denotes the vector with components $y^{(i)}(x_n) - y_n^{(i)}$. I is the identity matrix and G is the Jacobian matrix with elements $G_{ij} = \partial F_i / \partial y^{(j)}$ which are assumed constant, as in the case of a single equation. A cursory analysis of numerical stability is available through consideration of a characteristic polynomial corresponding to a majorization of equation (15). However, a more detailed approach involving the eigenvalues of the matrix G has been pursued in the present study.

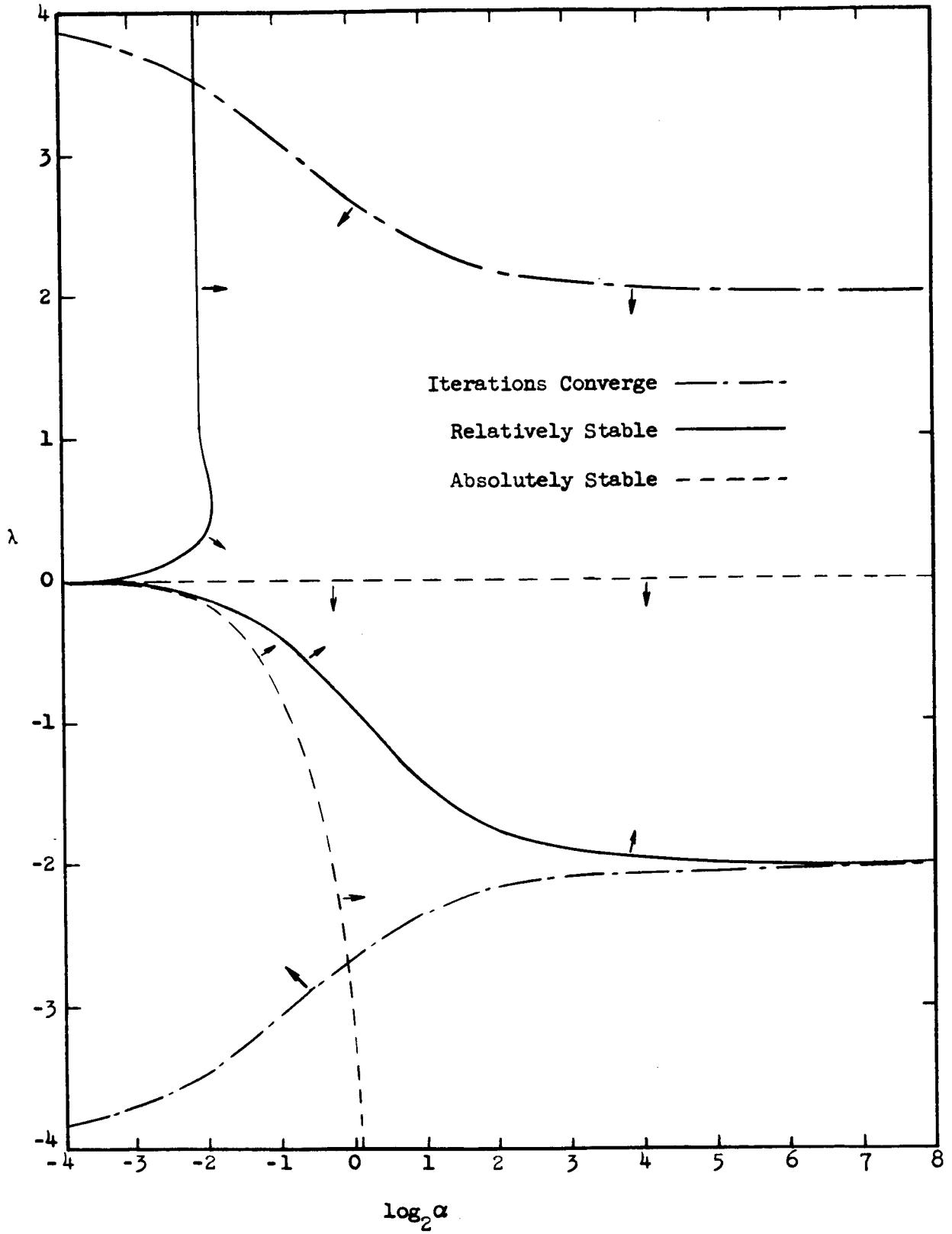

Figure 1. Regions of Stability and Convergence of Corrector Iterations of Variable Mesh Method

Pre-multiplying equation (15) by a matrix T, representing a nonsingular linear transformation such that $TGT^{-1} = J$ is in cononical form, gives

$$(I-d_{-1}hJ)\bar{\eta}_{n+1} - (I+d_0hJ)\bar{\eta}_n - d_1hJ\bar{\eta}_{n-1} - d_2hJ\bar{\eta}_{n-2} = 0, \qquad (16)$$

where $\bar{\eta}_n = T\bar{\epsilon}_n$. The diagonal elements of J are the eigenvalues of G, and if these are distinct, all the off diagonal elements of J are zero. In this case the system of difference equations for the propagated errors becomes uncoupled in passing from (15) to (16), and the relevant characteristic polynomial equation is again given by (13), with λ taking on the values hJ_{ii}. If the eigenvalues of G are not distinct, the analysis is more complicated, as indicated in [1], but the results are essentially the same. In either case, however, Fig. 1 is inadequate because some of the J_{ii} may have nonzero imaginary parts.

It is easy to show that the zeros of any polynomial whose coefficients are themselves polynomials in a complex variable λ are the complex conjugates of the zeros of the same polynomial with λ replaced by its conjugate. Thus we need only track the roots of (13) for values of λ with positive imaginary parts, the regions of numerical stability in the lower half of the λ-plane then being given by symmetry.

The problem of determining regions of stability for fixed α has thus been reduced to computing the roots of (13) for incremental values of λ in the upper half λ-plane and deciding at each point whether or not we have stability according to some appropriate definition involving the roots. We will limit ourselves to relative stability.

Choosing a definition of relative numerical stability presents an interesting situation. (We ignored this situation in the case of a single

differential equation. It was present but rather inconsequential.) One would like a definition which not only provides a unique decision regarding stability at each point but also reflects one's intuitive notions of relative stability. For example, it is distressing to find it possible to pass repeatedly back and forth from stability to instability as $|\lambda|$ increases along some specified path. Two definitions were considered in the present study--one an extension of the Ralston definition used above for single differential equations, and the other a definition used by Crane and Klopfenstein [6] and also by Krogh [8]. Both definitions lead to meaningless relative stability boundaries for fairly large complex λ. As a practical matter, however, it should be remembered that numerical stability is irrelevant for sufficiently large λ since either the truncation error becomes prohibitively large or convergence of the corrector iterations is not obtained.

The generalization of Ralston's definition to apply to systems was considered by Lea [9]. Lea defined the principal root of the characteristic polynomial equation as the continuous function of h satisfying the polynomial equation and taking on the value unity at h = 0. All others were called extraneous. Actually however, this "definition" fails to distinguish between the principal and extraneous roots because two of them may satisfy the requirements of the principal root. The following example illustrates this deficiency and further illustrates the inability to decide between stability and instability for a particular value of λ.

For $\alpha \equiv 1$ (fixed step size) the three roots of equation (13) are shown in the ρ-plane (Fig. 2). The values corresponding to $\lambda = (-1,2)$ are indicated by circles. Moving from the origin in the λ-plane counterclockwise around the rectangle to $(-1,2)$, the roots proceed in the ρ-plane from the points $(1,0)$, $(0,0)$, and $(0,0)$ to the circled points along the

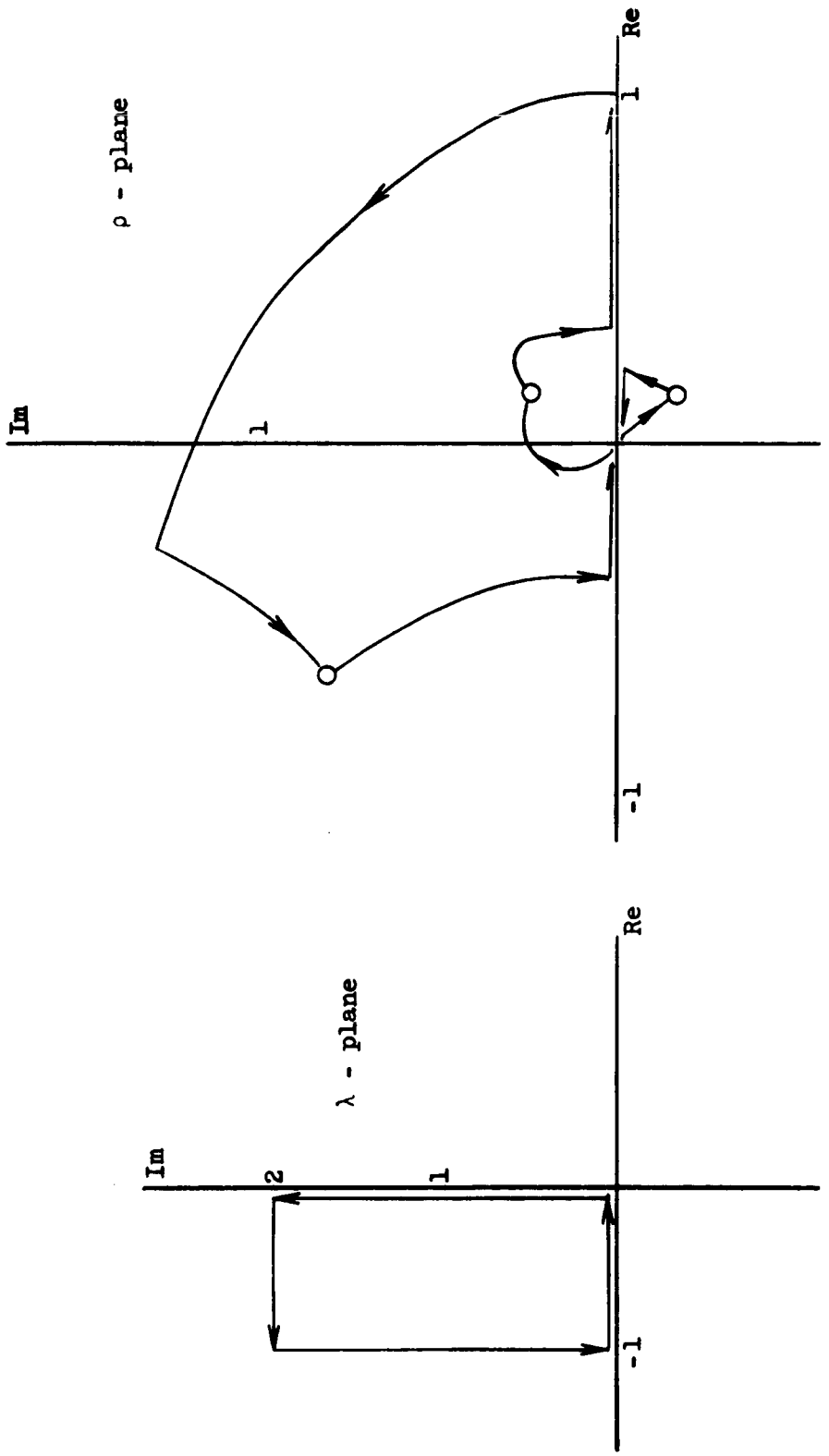

Figure 2. Example of Ambiguity in Relative Stability Definition

paths indicated by the arrows. The point $\lambda = (-1,2)$ appears stable according to the Lea definition since the root which started at $(1,0)$ is the largest. However, as we continue around the rectangle in λ-plane we see upon returning to the origin that the root which started at $(1,0)$ is now at $(0,0)$, while one of the roots which started at $(0,0)$ is now at $(1,0)$. In other words, if we had proceeded clockwise in the λ-plane, the point $\lambda = (-1,2)$ would appear unstable.

This problem does not develop with small λ; that is, when we consider a somewhat smaller rectangle the roots return to their starting points. On the other hand, the problem does preclude a complete partitioning of the λ-plane into meaningful regions of stability and instability by this procedure.

The alternate definition does uniquely partition the λ-plane into regions of stability and instability, but these regions are not acceptable for large λ . The problem here, although not recognized in either [6] or [8], is the one mentioned earlier of alternating between stability and instability as λ increases. According to this definition, a method is relatively stable if the modulus of each of the roots, other than the one nearest $\exp(\lambda)$, is less than or equal to $\exp[\mathrm{Re}(\lambda)]$, with equality permitted for simple roots only.

To illustrate the problem with this definition we note first that for $\alpha = 1$, the roots of equation (13) go from the "source points," $(1,0)$, $(0,0)$ and $(0,0)$, to the "sink points," approximately $(-2.37,0.0)$, $(0.13,-0.17)$, and $(0.13,+0.17)$, not necessarily respectively, as λ goes from the origin to infinity along any path in the λ-plane. Consider now, for example, λ moving along the real axis to $(0.5,0.0)$ and then vertically to infinity. For the vertical portion, $\exp(\lambda)$ traverses again and again

the circle in the ρ-plane with radius $\exp(0.5)$ and center $(0,0)$. Eventually, when the three roots are sufficiently close to their sink points, they are each nearest $\exp(\lambda)$ for a portion of each cycle of $\exp(\lambda)$. Thus by definition the method is relatively stable for the portion of each cycle when the root near $(-2.37,0.0)$ is the closest to $\exp(\lambda)$ and unstable otherwise. In this manner, on the vertical line $\mathrm{Re}(\lambda) = 0.5$, we have stability up to $\mathrm{Im}(\lambda) = 3.0$, then instability to about 8.3, stability again to about 9.9, etc.

Since the second definition has the practical advantage that its application is independent of path in the λ-plane, and since the problem just noted apparently occurs only for excessively large λ, there is no practical difficulty in its usage: one simply ignores stable regions lying "outside" unstable regions.

Consequently the results shown in Fig. 3 were obtained by applying the second definition. The two definitions give very similar results for small λ and reasonable values of α, say $1/4 < \alpha < 4$.

Also shown in Fig. 3 are the curves $\left|\lambda d_{-1}\right| = 1$. In a manner analogous to the case of a single differential equation, it can be shown that for the dominating eigenvalue of the Jacobian matrix of the system, the condition $\left|\lambda d_{-1}\right| < 1$ is necessary for convergence of the corrector iterations.

4. CRITERIA FOR SELECTING MESH INCREMENTS. An algorithm for the solution of differential equations by variable mesh procedures would be incomplete without reasonably sound, general purpose criteria for deciding what step size to use at each step of the integration. The main information required for specifying effective criteria was developed in the previous sections. In essence, the mesh selection procedure discussed below

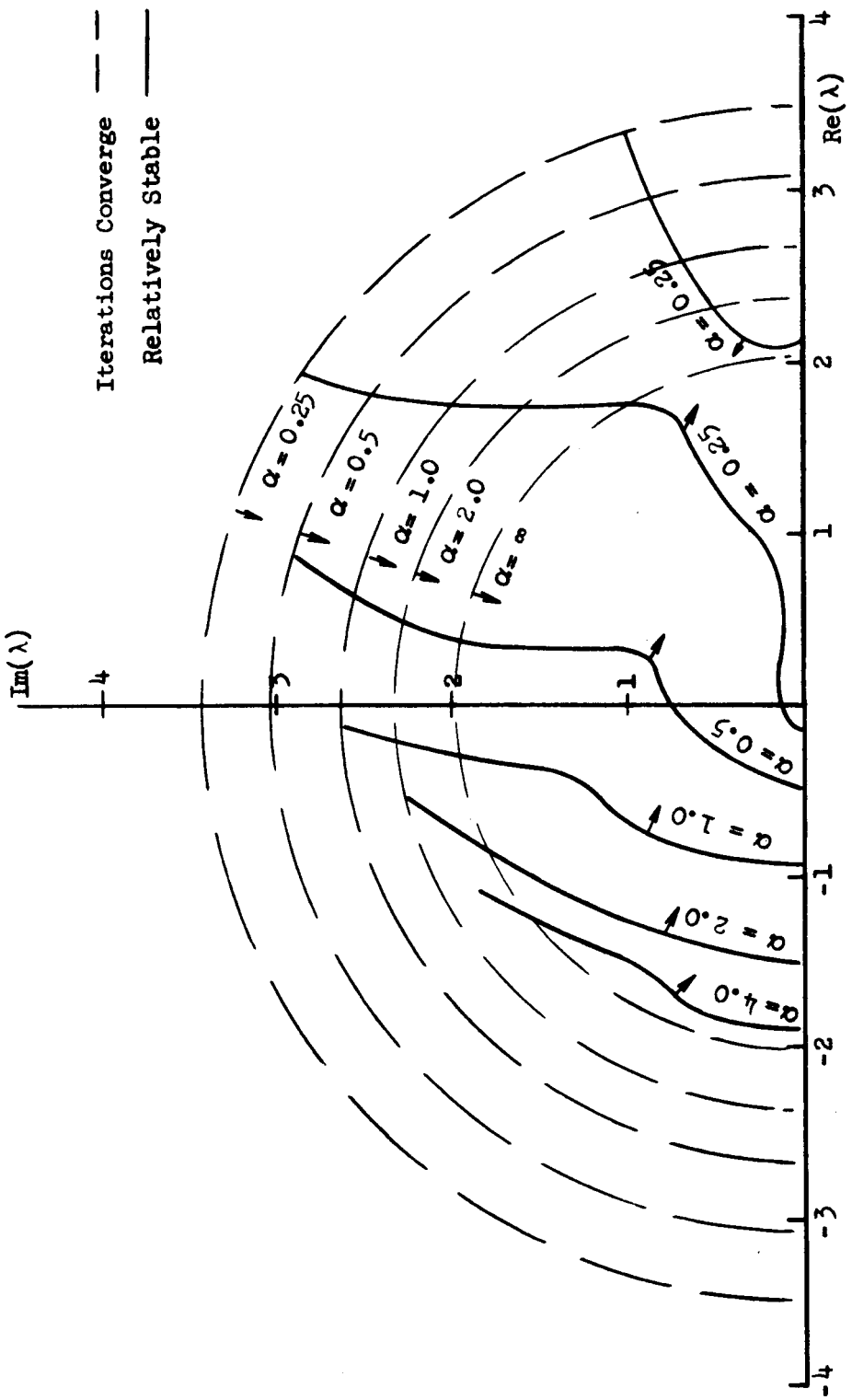

Figure 3. Regions of Relative Stability and Convergence of Corrector Iterations of Variable Mesh Method for Systems of Differential Equations

represents an attempt to choose each step size just small enough so that the following three criteria are satisfied:

a) The relative truncation error must remain within a prescribed tolerance δ.

b) The condition for convergence of the corrector iterations must be satisfied.

c) The method must possess relative numerical stability.

Let p_{n+1} and c_{n+1} denote the predicted and final corrected approximations for $y(x_{n+1})$. Let $H = x_{n+2} - x_{n+1}$ be the step size to be used in computing the solution at x_{n+2}, and let α_t be the new value of α as determined by the truncation error criterion in a manner described below. (Thus from the truncation error criterion we will get $H = h/\alpha_t$.)

Using the truncation error terms for the predictor and corrector formulas obtained in Section 2, we can eliminate the factor $\frac{h^5}{5!} y_n^v$ and obtain, through fifth order in h, the equation

$$y(x_{n+1}) - c_{n+1} = \left(\frac{c_{n+1} - p_{n+1}}{P_n - C_n}\right) C_n,$$

where P_n and C_n are given by equations (9) and (10), respectively.

We want to find α_t such that the relative error in c_{n+2} is δ, that is,

$$\left| y(x_{n+2}) - c_{n+2} \right| = \delta \left| y(x_{n+2}) \right|.$$

In practice we actually set

$$\left| \left(\frac{c_{n+1} - p_{n+1}}{P_n - C_n}\right) \frac{C_n}{h^5} \right| H^5 = \delta \left| c_{n+1} \right|,$$

giving

$$\alpha_t = \left| \frac{C_n (c_{n+1} - p_{n+1})}{\delta c_{n+1} (P_n - C_n)} \right|^{1/5} . \qquad (17)$$

Criteria b and c above are combined to produce a single value α_c for the mesh parameter α at the new step. To this end, we solve $H = h/\alpha_c$ simultaneously with expressions approximating the boundary of the intersection of the regions of relative stability and iteration convergence shown in Fig. 1. For this purpose the following expressions have been found to fit the boundary data accurately:

$$f_y < 0: \quad 0 < \alpha_c < .25: \quad Hf_y = -3.2\alpha_c^{5/2}$$

$$.25 \leq \alpha_c < 1.0: \quad Hf_y = .17 - 1.09\alpha_c$$

$$1.0 \leq \alpha_c < \infty: \quad Hf_y = 1.08/\alpha_c - 2,$$

$$f_y > 0: \quad \alpha_c < .25: \quad \text{not permitted (see Fig. 1)}$$

$$.25 \leq \alpha_c \leq 1.0: \quad Hf_y = \frac{4}{3}\left[2 + (1 - \alpha_c)^{7/4}\right]$$

$$1.0 < \alpha_c < \infty: \quad Hf_y = 2 + \frac{2}{3\alpha_c} .$$

An approximation for f_y can be obtained from computations from the completed step:

$$f_y \approx \frac{f(x_{n+1}, p_{n+1}) - f(x_{n+1}, c_{n+1})}{p_{n+1} - c_{n+1}} .$$

For the case $f_y > 0$, $.25 \leq \alpha_c \leq 1.0$, an iterative scheme is used to solve for α_c:

$$\alpha_c^{(i+1)} = \frac{3hf_y}{8 + 4(1 - \alpha_c^{(i)})^{7/4}} .$$

It has been determined (by actual calculations) that with $\alpha_c^{(0)} = hf_y/3$, $\alpha_c^{(1)}$ is always correct to within two units in the second decimal place. Thus α_c is computed according to the following simultaneous solutions of each of the above equations with the equation $H = h/\alpha_c$:

$$-\infty < hf_y \leq -.92: \quad \alpha_c = (1.08 - hf_y)/2$$

$$-.92 < hf_y \leq -.025: \quad \alpha_c = (.17 + \sqrt{.03 - 4.36hf_y})/2.18$$

$$-.025 < hf_y < 0: \quad \alpha_c = (-hf_y/3.2)^{2/7}$$

$$0 \leq hf_y \leq .875: \quad \alpha_c = .25$$

$$.875 < hf_y < 8/3: \quad \alpha_c = 3hf_y/\left[8 + 4(1 - hf_y/3)^{7/4}\right]$$

$$8/3 \leq hf_y < \infty: \quad \alpha_c = (hf_y - 2/3)/2.$$

The new step size H is then h/α, where $\alpha = \max(\alpha_t, \alpha_c)$.

It is the writer's experience that when even only a moderate degree of accuracy is required, the numerical solution of most problems is limited by accuracy rather than stability (or convergence of the iterations). This is because (relative) numerical stability requires only that the computed solution follow the primary trend in the true solution. If steps as large as possible for stability are taken, the secondary trends will not be evident, and the computed solution may be grossly inaccurate. When

we begin to examine mesh criteria for large systems of differential equations, it is especially fortuitous that satisfying the truncation error criterion usually precludes instability because in this case the truncation criterion is the only one which can be feasibly incorporated into the algorithm. For large systems the amount of computing time required to evaluate either the Jacobian matrix G or its eigenvalues at each step would usually be prohibitive. Of course for certain small systems it may not be prohibitive, and then the results shown in Fig. 3 can be incorporated in a manner analogous to that given above for obtaining α_c in the case of a single differential equation. This procedure has proved successful for selected systems although it did not alter the mesh increments substantially from those selected by the truncation criterion alone when reasonably small values of δ were used in the latter criterion.

The mesh selection procedure recommended for most large systems thus consists of using only the truncation error criterion. Values of α_t^5 are computed from equation (17) for each component of the system, and then α is set equal to the fifth root of the largest of these.

5. NUMERICAL TESTING AND COMPARISON WITH OTHER METHODS. The variable mesh multistep method has been tested by applying it to several single differential equations and to several systems of differential equations. This testing has given a fairly thorough demonstration of the effectiveness and reliability of the algorithm. One system of substantial importance for which the variable mesh approach proved especially effective was the problem of heat transfer to a supercritical fluid with variable physical properties and fully developed turbulent flow in a smooth tube [1]. Another system, discussed in Ref. 10, was a stochastic model of enzymatically controlled cooperative unwinding and template replication of biological macromolecules. Due to the complicated

mathematical formulation of these problems, they will not be given in detail here. However, several simpler test problems are listed in Table 1.

Most of these test problems were selected because of their inherent potential, both in the behavior of the solutions and in the behavior of the partial derivatives of the right hand sides with respect to the dependent variables, for producing numerical difficulties. Some are particularly suited to a variable mesh treatment while others, No. 5, 6, 10 and 11, can be solved efficiently with constant mesh increments. In the latter cases it is important to note that the accuracy obtained by the variable mesh method was about the same as that obtained using constant increments with the same number of steps. This indicates that the variable mesh procedures do not have a degrading effect when they are used unnecessarily.

Each equation was solved on the IBM System 360 using single precision starting values and double precision arithmetic to advance the solution. Values of δ, the target relative truncation error, ranging from 10^{-6} to 10^{-1} were used for each equation. The accuracy obtained was roughly proportionate to the values of δ specified. It was noted that the step lengths were limited almost entirely by the truncation error for the smaller values of δ with the stability/convergence criterion becoming of increasing importance with increasing δ.

Some of these problems, 1, 9, and 12, were used in comparing the new algorithm with other fourth order numerical methods which also permit some variability in the mesh increments. The other methods used were the standard fourth order Runge-Kutta method, the Nordsieck method, and the basic Adams-Bashforth/Adams-Moulton method, allowing doubling and halving of the increments with the latter. As indicated below, the new method proved superior to the other methods for these problems.

TABLE 1

PROBLEMS USED TO TEST VARIABLE MESH METHOD

Problem	Differential Equation(s)	Integration Interval	Initial Value	Exact Solution
1	$y' = -40xy$	$-1 \le x \le 1$	$\exp(-10)$	$\exp(10-20x^2)$
2	$y' = (2xy)^{-1}$	$1 \le x \le 10^{20}$	0	$\sqrt{\ln(x)}$
3	$y' = y/x - (1/x)\cos(1/x)$	$-1 \le x \le -.01$	$\sin(1)$	$x \sin(1/x)$
4	$y' = -\exp(x)\, y$	$0 \le x \le 5$	$\exp(-1)$	$\exp[-\exp(x)]$
5	$y' = -y$	$0 \le x \le 10$	1	$\exp(-x)$
6	$y' = y$	$0 \le x \le 10$	1	$\exp(x)$
7	$y' = -y/z$ $z' = -z$	$0 \le x \le 5$	$\exp(-1)$ 1	$\exp[-\exp(x)]$ $\exp(-x)$
8	$y' = y(y/z+1)$ $z' = y$	$0 \le x \le 5$	$-\exp(-1)$ $\exp(-1)$	$-\exp[x-\exp(x)]$ $\exp[-\exp(x)]$
9	$y' = y^2/z-40z$ $z' = y$	$-1 \le x \le 1$	$40 \exp(-10)$ $\exp(-10)$	$-40x \exp(10-20x^2)$ $\exp(10-20x^2)$
10	$y' = -2(y+z)$ $z' = y$	$0 \le x \le 100$	0 1	$-2 \exp(-x) \sin(x)$ $\exp(-x)[\sin(x) + \cos(x)]$
11	$y' = -\exp(-x)-100z$ $z' = -100\, z$	$0 \le x \le 1.5$	2 1	$\exp(-x)+ \exp(-100x)$ $\exp(-100x)$
12	$y' = -z/x^4$ $z' = y$	$-1 \le x \le -.01$	$\cos(1)-\sin(1)$ $\sin(1)$	$\sin(1/x)-(1/x)\cos(1/x)$ $x \sin(1/x)$

Since the Runge-Kutta method requires four derivative evaluations per step while the others were used with only two evaluations per step, half as many steps were used with the Runge-Kutta method as with the other three.

The Nordsieck method permits increasing (or decreasing) the step size by a factor θ (or $1/\theta$). The test problems used in the present study were solved with $\theta = 2$, the value emphasized in [4] where the symbol "β" is used for this factor, and also with smaller values to permit more gradual varying of the increments. In addition, Nordsieck's interval control mechanism requires a parameter "e" used in a manner to imply a target error θ^{-e}. For each value of θ, the problems used here were solved with several values of e, seeking one which produced the number of steps commensurate with the number used by the other methods. However for $\theta = 2$, the Nordsieck method used too many steps even when e was reduced to unity. (In fact, considerable difficulty was encountered in trying to locate values of θ which were usable in this sense. Successful choices are indicated in Table 2.) It is also noted here that it was not necessary to use Nordsieck's starting procedure for the test problems since all the required initial information was available.

For Problem 1, the absolute value of the relative error in the solution obtained by each of the four methods is shown in Fig. 4. For this problem, the entries in Table 2 are the areas under the curves of Fig. 4. For the other two problems, the entries in Table 2 reflect alternative measures of relative error which are more appropriate for the numerical solution obtained for those two systems of equations. As can be seen from the table, the new variable mesh method gave the best performance; and the basic Adams method, augmented with interpolation procedures to permit doubling and halving, also did considerably better than the other two methods.

TABLE 2

COMPARISON OF RELATIVE ERROR

Problem	Variable Mesh	Adams	Runge-Kutta	Nordsieck
1	1.5×10^{-4}	2.3×10^{-4}	2.8×10^{-3}	3.0×10^{-3} ($\theta = 1.01$)
9	1.8×10^{-3}	6.0×10^{-3}	5.2×10^{-2}	1.7×10^{-2} ($\theta = 1.01$)
12	2.0×10^{-2}	5.5×10^{-2}	1.0×10^{-1}	6.5×10^{-1} ($\theta = 1.5$)

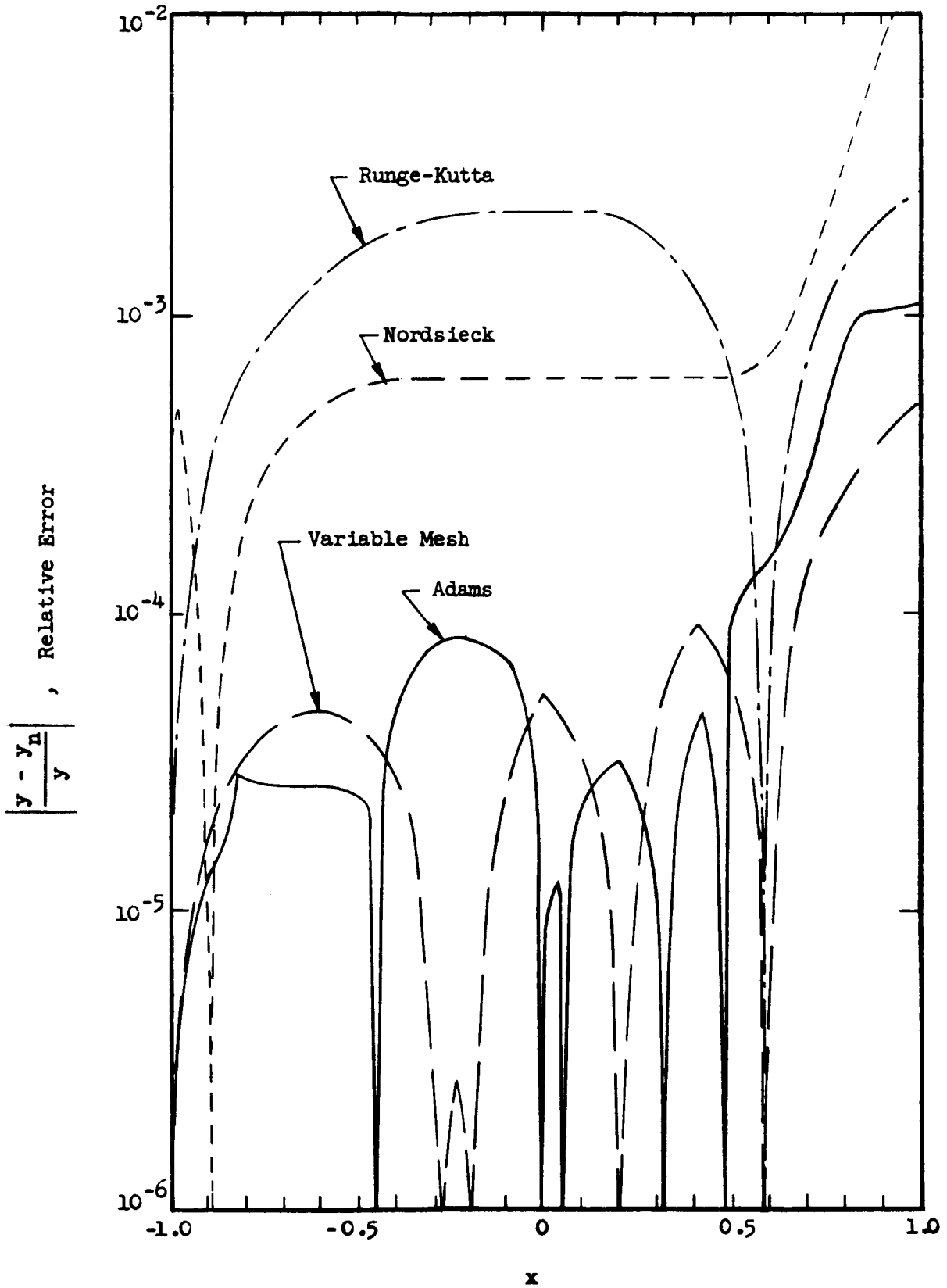

Figure 4. Relative Errors for Problem 1

REFERENCES

1. Van Wyk, R. A variable mesh method for the efficient numerical solution of ordinary differential equations. Rocketdyne Research Report No. 65-26, 1965, presented at the 1966 National SIAM meeting, Iowa City, Iowa.

2. Treanor, C. E. A method for the numerical integration of coupled first-order differential equations with greatly different time constants. Math. Comp., 20 (1966).

3. Henrici, P. Discrete Variable Methods in Ordinary Differential Equations. John Wiley and Sons, Inc., New York, 1962.

4. Nordsieck, A. On numerical integration of ordinary differential equations. Math. Comp., 16 (1962).

5. Hull, T. E. and Creemer, A. L. Efficiency of predictor-corrector procedures. J. ACM, 10 (1963)

6. Crane, R. L. and Klopfenstein, R. W. A predictor-corrector algorithm with an increased range of absolute stability. J. ACM, 12 (1965).

7. Ralston, A. Relative stability in the numerical solution of ordinary differential equations. SIAM Rev., 7 (1965).

8. Krogh, F. T. Predictor-corrector methods of high order with improved stability characteristics. J. ACM, 13 (1966).

9. Lea, R. N. On the stability of numerical solutions of ordinary differential equations. NASA Technical Note No. D-3760, 1967.

10. Zimmerman, J. M. and Simha, R. The kinetics of cooperative unwinding and template replication of biological macromolecules. J. Theoret. Biol., 13 (1966).

UNCLASSIFIED

VARIABLE MESH MULTISTEP METHODS FOR ORDINARY DIFFERENTIAL EQUATIONS

R. Van Wyk
Research Report 67-12, September 1967

1. Variable Mesh Methods
2. Numerical Analysis
3. Ordinary Differential Equations

UNCLASSIFIED

VARIABLE MESH MULTISTEP METHODS FOR ORDINARY DIFFERENTIAL EQUATIONS

R. Van Wyk
Research Report 67-12, September 1967

1. Variable Mesh Methods
2. Numerical Analysis
3. Ordinary Differential Equations

UNCLASSIFIED

VARIABLE MESH MULTISTEP METHODS FOR ORDINARY DIFFERENTIAL EQUATIONS

R. Van Wyk
Research Report 67-12, September 1967

1. Variable Mesh Methods
2. Numerical Analysis
3. Ordinary Differential Equations

UNCLASSIFIED

VARIABLE MESH MULTISTEP METHODS FOR ORDINARY DIFFERENTIAL EQUATIONS

R. Van Wyk
Research Report 67-12, September 1967

1. Variable Mesh Methods
2. Numerical Analysis
3. Ordinary Differential Equations

UNCLASSIFIED ABSTRACT

A multistep predictor-corrector method for the numerical solution of ordinary differential equations is developed. The difference equations employed are generalizations, for the case of variable mesh spacing, of previous formulas requiring fixed step size. In addition to retaining the high local accuracy and convergence properties of the earlier methods, the variable mesh method is developed in a form conducive to the generation of effective criteria for the selection of subsequent step sizes in the step by step solution of differential equations. These criteria are based on considerations of truncation error, convergence of corrector iterations, and an extensive treatment of relative numerical stability. The algorithm has been tested extensively and compared with other methods. The results of the comparison favor the new method.

UNCLASSIFIED ABSTRACT

A multistep predictor-corrector method for the numerical solution of ordinary differential equations is developed. The difference equations employed are generalizations, for the case of variable mesh spacing, of previous formulas requiring fixed step size. In addition to retaining the high local accuracy and convergence properties of the earlier methods, the variable mesh method is developed in a form conducive to the generation of effective criteria for the selection of subsequent step sizes in the step by step solution of differential equations. These criteria are based on considerations of truncation error, convergence of corrector iterations, and an extensive treatment of relative numerical stability. The algorithm has been tested extensively and compared with other methods. The results of the comparison favor the new method.

UNCLASSIFIED ABSTRACT

A multistep predictor-corrector method for the numerical solution of ordinary differential equations is developed. The difference equations employed are generalizations, for the case of variable mesh spacing, of previous formulas requiring fixed step size. In addition to retaining the high local accuracy and convergence properties of the earlier methods, the variable mesh method is developed in a form conducive to the generation of effective criteria for the selection of subsequent step sizes in the step by step solution of differential equations. These criteria are based on considerations of truncation error, convergence of corrector iterations, and an extensive treatment of relative numerical stability. The algorithm has been tested extensively and compared with other methods. The results of the comparison favor the new method.

UNCLASSIFIED ABSTRACT

A multistep predictor-corrector method for the numerical solution of ordinary differential equations is developed. The difference equations employed are generalizations, for the case of variable mesh spacing, of previous formulas requiring fixed step size. In addition to retaining the high local accuracy and convergence properties of the earlier methods, the variable mesh method is developed in a form conducive to the generation of effective criteria for the selection of subsequent step sizes in the step by step solution of differential equations. These criteria are based on considerations of truncation error, convergence of corrector iterations, and an extensive treatment of relative numerical stability. The algorithm has been tested extensively and compared with other methods. The results of the comparison favor the new method.